

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 09:14:43 ; Search time 73 Seconds
(without alignments)
5027.885 Million cell updates/sec

Title: US-09-671-687A-3
Perfect score: 949
Sequence: 1 MSSGLWSQKVTSPTWEEIR.....RLLCDAYMCYQPTMSLYK 949

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980a:*
- 2: Geneseqp1990a:*
- 3: Geneseqp2000a:*
- 4: Geneseqp2001a:*
- 5: Geneseqp2002a:*
- 6: Geneseqp2003a:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	606	63.9	685	4	Aam39254 Human pol
2	606	63.9	731	3	Aay91954 Human cyt
3	606	63.9	739	4	Aau23747 Novel hum
4	606	63.9	953	7	Adc24816 Human bre
5	606	63.9	953	8	Adg95918 T cell ac
6	606	63.9	956	6	Abb82783 Human CYL
7	606	63.9	956	8	Adr14489 Human NF-
8	606	63.9	960	8	Adi-99244 DKF2p586D
9	505	53.2	731	4	Aab95828 Human pro
10	505	53.2	953	8	Adg95916 T cell ac
11	505	53.2	956	8	Adg95920 T cell ac
12	363	38.3	698	4	Aam41040 Human pol
13	330	34.8	476	4	Aab95719 Human pro
14	227	23.9	558	4	Aau23211 Novel hum
15	208	21.9	261	5	Abb89233 Human pol
16	148	15.6	238	5	Abb89234 Human pol
17	113	11.9	113	4	Aam14965 Peptide #
18	113	11.9	113	4	Abb33941 Peptide #
19	113	11.9	113	4	Aam27399 Peptide #
20	113	11.9	113	4	Abb28754 Peptide #
21	113	11.9	113	4	Abb19377 Protein #
22	113	11.9	113	4	Aam67104 Human bon
23	113	11.9	113	4	Aam54704 Human bra
24	113	11.9	113	4	Aam02691 Peptide #
25	113	11.9	113	5	Abg36764 Human pep

26	106	11.2	106	8	ADK71947	Adk71947 Human ori
27	101	10.6	101	8	ADK71945	Adk71945 Human ori
28	96	10.1	96	8	ADK71949	Adk71949 Human ori
29	91	9.6	91	8	ADK71943	Adk71943 Human ori
30	91	9.6	104	8	ADK71962	Adk71962 Human ori
31	10	1.1	517	4	ABB61669	Abb61669 Drosophill
32	8	0.8	78	5	ABP68782	Abp68782 Marine sn
33	8	0.8	120	5	ADK35501	Adk35501 Novel hum
34	8	0.8	216	4	AAg72831	AAg72831 Human olf
35	8	0.8	228	7	ABO73471	ABO73471 Pseudomon
36	8	0.8	418	4	ABB64546	Abb64546 Drosophill
37	8	0.8	450	5	ABB50103	Abb50103 Listeria
38	8	0.8	450	6	ABU32881	Abu32881 Proteina
39	8	0.8	551	6	ABR83636	ABr83636 Saccharom
40	8	0.8	644	7	ADB80358	ADB80358 Human MDD
41	8	0.8	708	7	ABO71275	ABO71275 Pseudomon
42	8	0.8	941	7	ADF76391	Adf76391 Novel hum
43	8	0.8	1004	7	ADB80336	ADB80336 Human MDD
44	8	0.8	1086	8	ADN23310	Adn23310 Bacterial
45	7	0.7	18	5	AAO15840	AAO15840 CtxB sign

ALIGNMENTS

RESULT 1
AAM39254
ID AAM39254 standard; protein; 685 AA.

XX AC AAM39254;

XX AC AAM39254;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 2399.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00662191.

XX 19-OCT-2000; 2000US-00693036.

XX 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren P, Wang D;
Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
Zhou P, Goodrich R, Drmanac RT;
WPI; 2001-442253/47.
XX N-PSDB; AAI58410.
XX Novel nucleic acids and polypeptides, useful for treating disorders such
as central nervous system injuries.
XX Example 4; SEQ ID NO 2399; 10078pp; English.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 09:18:24 ; Search time 57 seconds
(without alignments)
1242.841 Million cell updates/sec

Title: US-09-671-687A-3
Perfect score: 949
Sequence: 1 MSSGLWSQKVTSPYWEERI.....RLLCDAYMCWQSPMSLYK 949

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/ECTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	0.8	228	4	US-09-252-991A-22217
2	8	0.8	344	4	US-09-270-767-42367
3	8	0.8	708	4	US-09-252-991A-20021
4	8	0.8	2216	4	US-09-902-540-12221
5	7	0.7	28	3	US-09-227-357-607
6	7	0.7	56	4	US-09-513-999C-6825
7	7	0.7	62	1	US-08-616-732A-24
8	7	0.7	62	3	US-09-037-742B-24
9	7	0.7	108	4	US-09-513-999C-6103
10	7	0.7	112	4	US-09-902-540-14754
11	7	0.7	116	2	US-08-879-995A-4
12	7	0.7	116	3	US-09-215-096-4
13	7	0.7	125	4	US-09-583-110-4643
14	7	0.7	127	4	US-09-513-999C-4224
15	7	0.7	134	4	US-09-107-433-2866
16	7	0.7	135	4	US-09-270-767-36133
17	7	0.7	135	4	US-09-270-767-51350
18	7	0.7	140	4	US-09-270-767-32498
19	7	0.7	140	4	US-09-270-767-47715
20	7	0.7	142	4	US-09-602-777A-334
21	7	0.7	144	4	US-09-248-796A-18078
22	7	0.7	147	1	US-08-688-609-8
23	7	0.7	147	3	US-09-002-832-8
24	7	0.7	151	4	US-09-270-767-44667
25	7	0.7	151	4	US-09-902-540-15459
26	7	0.7	152	4	US-09-621-976-4852
27	7	0.7	153	4	US-09-270-767-56889

28	7	0.7	156	4	US-09-710-279-1222	Sequence 1222, Ap
29	7	0.7	162	4	US-09-270-767-31034	Sequence 31034, A
30	7	0.7	162	4	US-09-270-767-36622	Sequence 36622, A
31	7	0.7	162	4	US-09-270-767-48251	Sequence 48251, A
32	7	0.7	162	4	US-09-270-767-51839	Sequence 51839, A
33	7	0.7	163	2	US-08-783-395-5	Sequence 5, Appl1
34	7	0.7	163	3	US-08-600-430-2	Sequence 2, Appl1
35	7	0.7	163	3	US-09-134-001C-4131	Sequence 4131, Ap
36	7	0.7	163	4	US-09-461-912A-44	Sequence 44, Appl
37	7	0.7	163	4	US-09-949-016-6077	Sequence 6077, Ap
38	7	0.7	169	4	US-09-711-164-381	Sequence 381, App
39	7	0.7	172	4	US-09-270-767-41649	Sequence 41649, A
40	7	0.7	174	4	US-09-949-016-11208	Sequence 11208, A
41	7	0.7	181	4	US-09-513-999C-5797	Sequence 5797, Ap
42	7	0.7	185	4	US-09-583-110-4438	Sequence 4438, Ap
43	7	0.7	190	4	US-09-328-352-6627	Sequence 6627, Ap
44	7	0.7	192	4	US-09-107-433-3944	Sequence 3944, Ap
45	7	0.7	193	4	US-09-252-991A-17828	Sequence 17828, A

ALIGNMENTS

RESULT 1
US-09-252-991A-22217
; Sequence 22217, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22217
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22217

Query Match 0.8%; Score 8; DB 4; Length 228;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 VELLEGR 171
DB 151 VELLEGR 158

RESULT 2

US-09-270-767-42367
; Sequence 42367, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42367
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-42367

Query Match 0.8%; Score 8; DB 4; Length 344;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 09:40:40 ; Search time 57 Seconds
(without alignments)
5533.680 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 949

Sequence: 1 MSSGLWSQBKVTSPYWEERI.....RLLCDAYMCWQSTMSLYK 949

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1421835 seqs, 332370683 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	606	63.9	731	16 US-10-921-707-9	Sequence 9, Appli
2	606	63.9	956	10 US-09-851-673-4	Sequence 4, Appli
3	606	63.9	956	16 US-10-755-889-490	Sequence 490, App
4	606	63.9	960	16 US-10-788-792-250	Sequence 250, App
5	208	21.9	261	15 US-10-264-237-1609	Sequence 1609, Ap
6	148	15.6	238	15 US-10-264-237-1610	Sequence 1610, Ap
7	113	11.9	113	9 US-09-864-761-34675	Sequence 34675, A
8	0.8	0.8	49	15 US-10-424-599-186263	Sequence 186263,
9	0.8	0.8	78	14 US-10-058-053A-218	Sequence 218, App
10	0.8	0.8	78	16 US-10-838-226-218	Sequence 218, App
11	0.8	0.8	122	16 US-10-437-963-180681	Sequence 180681,
12	0.8	0.8	143	15 US-10-424-599-170638	Sequence 170638,
13	0.8	0.8	188	15 US-10-424-599-159539	Sequence 159539,

14	0.8	0.8	371	16 US-10-437-963-188186	Sequence 188186,
15	0.8	0.8	450	15 US-10-282-132A-60805	Sequence 60805, A
16	0.8	0.8	1086	15 US-10-369-493-5963	Sequence 5963, Ap
17	0.7	0.7	28	10 US-09-983-802-607	Sequence 607, App
18	0.7	0.7	28	10 US-09-983-802-607	Sequence 607, App
19	0.7	0.7	28	11 US-09-973-278-461	Sequence 461, App
20	0.7	0.7	30	16 US-10-632-706-150 ⁴	Sequence 150, App
21	0.7	0.7	30	16 US-10-632-706-154	Sequence 154, App
22	0.7	0.7	33	14 US-10-081-872-219	Sequence 219, App
23	0.7	0.7	33	15 US-10-385-305-219	Sequence 219, App
24	0.7	0.7	45	15 US-10-264-049-3820	Sequence 3820, Ap
25	0.7	0.7	53	15 US-10-613-413A-73	Sequence 73, Appl
26	0.7	0.7	53	17 US-10-885-225-73	Sequence 73, Appl
27	0.7	0.7	61	15 US-10-424-599-276896	Sequence 276896,
28	0.7	0.7	62	15 US-10-424-599-193894	Sequence 193894,
29	0.7	0.7	62	15 US-10-424-599-221399	Sequence 221399,
30	0.7	0.7	67	11 US-09-864-408A-6310	Sequence 6310, Ap
31	0.7	0.7	67	15 US-10-424-599-266715	Sequence 266715,
32	0.7	0.7	73	16 US-10-767-701-49557	Sequence 49557, A
33	0.7	0.7	79	16 US-10-437-963-184451	Sequence 184451,
34	0.7	0.7	82	16 US-10-767-701-61504	Sequence 61504, A
35	0.7	0.7	86	15 US-10-424-599-150683	Sequence 150683,
36	0.7	0.7	86	16 US-10-437-963-195228	Sequence 195228,
37	0.7	0.7	88	16 US-10-437-963-107983	Sequence 107983,
38	0.7	0.7	92	16 US-10-437-963-162456	Sequence 162456,
39	0.7	0.7	98	15 US-10-424-599-210307	Sequence 210307,
40	0.7	0.7	102	16 US-10-437-963-106579	Sequence 106579,
41	0.7	0.7	104	15 US-10-424-599-247009	Sequence 247009,
42	0.7	0.7	104	16 US-10-767-701-48578	Sequence 48578, A
43	0.7	0.7	106	11 US-09-864-408A-2374	Sequence 2374, Ap
44	0.7	0.7	106	16 US-10-437-963-199204	Sequence 199204,
45	0.7	0.7	107	16 US-10-663-244-83	Sequence 83, Appl

ALIGNMENTS

RESULT 1
US-10-921-707-9
; Sequence 9, Application US/10921707
; Publication No. US20050003447A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BANDMAN, Olga
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0594 PCT
; CURRENT APPLICATION NUMBER: US/10/921,707
; PRIOR FILING DATE: 2004-08-19
; PRIOR APPLICATION NUMBER: US/09/786,797
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: 09/156,470; unassigned; 60/131,321
; PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 2363327
US-10-921-707-9
Query Match 63.9%; Score 606; DB 16; Length 731;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 17, 2005, 01:26:12 ; Search time 8968 Seconds
(without alignments)

5127.565 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 949

Sequence: 1 MSGLWSQEKVTSFYWEERI.....RLLCDAYMCYQPTMSLYK 949

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 1

Total number of hits satisfying chosen parameters: 9408497

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlp
-Q=/cgn2.1/USPTO spool_p/US09671687/runat_15042005_163038_11950/app.query.fasta_1.1095
-DB=GenEm1 -QFT=fastbp -SUFFIX=oligo.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human4.0.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09671687 @CGN 1.1 5932 @runat_15042005_163038_11950 -NCPUs=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

GenEm1.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.ste.*
12: gb.ev.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	754	79.5	3302	6	CQ834224 Sequence
2	754	79.5	3540	9	BC012342 Homo sapi
3	754	79.5	5414	9	AB020656 Homo sapi
4	711	74.9	4527	6	BD231207 Human cyt

ALIGNMENTS

RESULT 1	CQ834224	CQ834224	Sequence 95 from Patent WO2004058805.	3302 bp	DNA	linear	PAT 29-JUL-2004
LOCUS	CQ834224	CQ834224	CQ834224	3302 bp	DNA	linear	PAT 29-JUL-2004
DEFINITION	CQ834224	CQ834224	CQ834224	3302 bp	DNA	linear	PAT 29-JUL-2004
ACCESSION	CQ834224	CQ834224	CQ834224	3302 bp	DNA	linear	PAT 29-JUL-2004
VERSION	CQ834224	CQ834224	CQ834224	3302 bp	DNA	linear	PAT 29-JUL-2004
KEYWORDS	CQ834224	CQ834224	CQ834224	3302 bp	DNA	linear	PAT 29-JUL-2004
SOURCE	CQ834224	CQ834224	CQ834224	3302 bp	DNA	linear	PAT 29-JUL-2004
ORGANISM	CQ834224	CQ834224	CQ834224	3302 bp	DNA	linear	PAT 29-JUL-2004
REFERENCE	CQ834224	CQ834224	CQ834224	3302 bp	DNA	linear	PAT 29-JUL-2004
AUTHORS	CQ834224	CQ834224	CQ834224	3302 bp	DNA	linear	PAT 29-JUL-2004
TITLE	CQ834224	CQ834224	CQ834224	3302 bp	DNA	linear	PAT 29-JUL-2004
JOURNAL	CQ834224	CQ834224	CQ834224	3302 bp	DNA	linear	PAT 29-JUL-2004
FEATURES	CQ834224	CQ834224	CQ834224	3302 bp	DNA	linear	PAT 29-JUL-2004
source	CQ834224	CQ834224	CQ834224	3302 bp	DNA	linear	PAT 29-JUL-2004
CDS	CQ834224	CQ834224	CQ834224	3302 bp	DNA	linear	PAT 29-JUL-2004

5	685	72.2	5371	9	HSA250014	AJ250014 Homo sapi
6	667	70.3	3302	6	CQ834222	CQ834222 Sequence
7	641	67.5	2523	6	AR338799	AR338799 Sequence
8	640	67.4	3480	6	CQ719792	CQ719792 Sequence
9	608	64.1	2845	6	BD160617	BD160617 Primer fo
10	608	64.1	2845	6	AX883937	AX883937 Sequence
11	608	64.1	2845	9	AK024348	AK024348 Homo sapi
12	598	63.0	3311	6	CQ834226	CQ834226 Sequence
13	517	54.5	2341	9	AK000187	AK000187 Homo sapi
14	418	44.0	2569	6	BD160470	BD160470 Primer fo
15	418	44.0	2569	6	AX883681	AX883681 Sequence
16	418	44.0	2569	9	AK024212	AK024212 Homo sapi
17	374	39.4	1954	9	AK056226	AK056226 Homo sapi
18	165	17.4	2116	6	BD135433	BD135433 Receptor
19	165	17.4	2116	6	CQ867781	CQ867781 Sequence
20	165	17.4	2116	6	AR534837	AR534837 Sequence
21	165	17.4	2116	6	AX017216	AX017216 Sequence
22	140	14.8	141663	2	AC145018	AC145018 Felis cat
23	140	14.8	155857	2	AC145174	AC145174 Pan trogl
24	140	14.8	163319	9	HSA303140	AJ303140 Homo sapi
25	140	14.8	168271	9	AC007728	AC007728 Homo sapi
26	140	14.8	178790	2	AC145238	AC145238 Pan trogl
27	124	13.1	837	6	BD149741	BD149741 Primer fo
28	124	13.1	837	6	AX869679	AX869679 Sequence
29	123	13.0	758	6	BD147076	BD147076 Primer fo
30	123	13.0	758	6	AX867014	AX867014 Sequence
31	116	12.2	3267	10	BC082001	BC082001 Rattus no
32	116	12.2	4314	10	AK122389	AK122389 Mus muscu
33	116	12.2	4501	10	BC042438	BC042438 Mus muscu
34	109	11.5	241990	2	AC123449	AC123449 Rattus no
35	109	11.5	251132	2	AC098162	AC098162 Rattus no
36	108	11.4	251957	2	AC126867	AC126867 Rattus no
37	107	11.3	2426	10	BC049879	BC049879 Mus muscu
38	102	10.7	309	6	CQ055561	CQ055561 Sequence
39	102	10.7	309	6	CQ074825	CQ074825 Sequence
40	102	10.7	309	6	CQ105734	CQ105734 Sequence
41	102	10.7	309	6	CQ144418	CQ144418 Sequence
42	102	10.7	309	6	CQ179895	CQ179895 Sequence
43	102	10.7	309	6	CQ204366	CQ204366 Sequence
44	102	10.7	309	6	CQ265785	CQ265785 Sequence
45	102	10.7	309	6	CQ302867	CQ302867 Sequence

Location/Qualifiers
1..3302
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
243..3104
/note="unnamed protein product"
/codon_start=1
/protein_id="CAH05327.1"
/db_xref="GI:50833762"
/translation="MSSGLWSQEKVTSFYWEERI FYLLLOECSTVDKOTKLLKVPKG
SIGVYIDRSVGHSRIPSAKGGKQKIGLKEQPHAVLPVDEKDVVEINEKTELLA

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 17, 2005, 01:25:02 ; Search time 1076. Seconds
(without alignments)
5221.036 Million cell updates/sec

Title: US-09-671-687A-3
Perfect score: 949
Sequence: 1 MSSGLWSQKVTSPYWEERI.....RLLCDAYMCWYQSPWLSLYK 949

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delopt 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 1

Total number of hits satisfying chosen parameters: 8776198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-O=/cn2.1/USPTO.spool.p/US09671687/runat.15042005.163037.11942/app.query.fasta.1.1095
-DB=N_Geneseq_16Dec04 -OFT=fastap -SUFFIX=oligo.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09671687 @CN 1.1 708 @runat.15042005.163037.11942 -NCPU=3
-NO MAP -LARGEQUERY -NEG SCORE=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:
1: Geneseqn1980s:
2: Geneseqn1990s:
3: Geneseqn2000s:
4: Geneseqn2001as:
5: Geneseqn2001bs:
6: Geneseqn2002as:
7: Geneseqn2002bs:
8: Geneseqn2003as:
9: Geneseqn2003bs:
10: Geneseqn2003cs:
11: Geneseqn2003ds:
12: Geneseqn2004as:
13: Geneseqn2004bs:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	754	79.5	3302	12	Adc95917 T cell ac
2	754	79.5	6331	10	Adc24898 Human bre
3	711	74.9	4527	3	Aaa08589 Human cyt
4	685	72.2	5371	10	Abv75394 Human CYL
5	685	72.2	5371	13	Adr14488 Human NF-

6	667	70.3	3302	12	ADQ95915	Adq95915 T cell ac
7	641	67.5	2523	4	AAI58410	AAI58410 Human pol
8	641	67.5	2523	5	ADQ98620	ADQ98620 DNA encod
9	641	67.5	2523	9	ADB48380	ADB48380 Novel hum
10	641	67.5	2523	9	AAI58410	AAI58410 Novel hum
11	608	64.1	2845	4	AAI18625	AAI18625 CDNA enco
12	598	63.0	3311	12	ADQ95919	ADQ95919 T cell ac
13	418	44.0	2569	4	AAH18478	AAH18478 Human CDN
14	402	42.4	4286	4	AAI60196	AAI60196 Human pol
15	279	29.4	1151	6	ABL89642	ABL89642 Human pol
16	262	27.6	2488	4	AAI1081	AAI1081 CDNA enco
17	171	18.0	617	6	ABQ60558	ABQ60558 Human col
18	165	17.4	2116	2	AAZ07515	AAZ07515 Human RAP
19	148	15.6	1013	6	ABL89643	ABL89643 Human pol
20	124	13.1	837	4	AAH07749	AAH07749 Human CDN
21	123	13.0	758	4	AAH05084	AAH05084 Human CDN
22	115	12.1	557	6	ABQ60559	ABQ60559 Human col
23	106	11.2	318	12	ADK71948	ADK71948 Human ori
24	102	10.7	309	4	AAI20692	AAI20692 Probe #10
25	102	10.7	309	4	ABA65741	ABA65741 Human foe
26	102	10.7	309	4	AAI45907	AAI45907 Probe #14
27	102	10.7	309	4	ABA47848	ABA47848 Human bre
28	102	10.7	309	4	ABA32825	ABA32825 Probe #11
29	102	10.7	309	4	AAK39883	AAK39883 Human bon
30	102	10.7	309	4	AAK14143	AAK14143 Human bra
31	102	10.7	309	5	AAI06390	AAI06390 Probe #63
32	102	10.7	309	6	ABS13981	ABS13981 Human gen
33	101	10.6	303	12	ADK71946	ADK71946 Human ori
34	98	10.3	483	4	AAI11479	AAI11479 Probe #14
35	98	10.3	483	4	ABA53149	ABA53149 Human foe
36	98	10.3	483	4	AAI32754	AAI32754 Probe #14
37	98	10.3	483	4	ABA42721	ABA42721 Human bre
38	98	10.3	483	4	ABA22920	ABA22920 Probe #13
39	98	10.3	483	4	AAK26850	AAK26850 Human bon
40	98	10.3	483	4	AAK01396	AAK01396 Human bra
41	98	10.3	483	5	AAI01393	AAI01393 Probe #13
42	98	10.3	483	6	ABS01449	ABS01449 Human gen
43	96	10.1	288	12	ADK71950	ADK71950 Human ori
44	91	9.6	273	12	ADK71944	ADK71944 Human ori
45	86	9.1	1160	5	AAI90524	AAI90524 DNA encod

ALIGNMENTS

RESULT 1
ADQ95917
ID ADQ95917 standard; CDNA; 3302 BP.
XX
AC ADQ95917;
XX
DT 07-OCT-2004 (first entry)
XX
DE T cell activation associated CDNA #48.
XX
KW ss; gene; anti-allergic; antiarthritic; antidiabetic;
KW anti-HIV; antimicrobial; antirheumatic; immunosuppressive;
KW neuroprotective; gene therapy; T cell activation; diagnosis;
KW autoimmune disease; rheumatoid arthritis; asthma; multiple sclerosis;
KW diabetes; allergic disease; infectious disease; AIDS; chronic rejection;
KW organ; bone-marrow transplant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 243..3104
FT /*tag= a
FT
XX
PN WO2004058805-A2.
XX
PD 15-JUL-2004.
XX
PP 25-DEC-2003; 2003WO-JP016715.
XX

Result No.	Score	Query Match	Length	DB	ID	Description
1	641	17.5	2523	4	US-09-620-312D-290	Sequence 290, App
2	165	17.4	2116	4	US-09-646-403-3	Sequence 3, Appli
3	36	3.8	110	4	US-09-513-999C-20085	Sequence 20085, A
C 4	9	0.9	657	3	US-08-998-416-248	Sequence 246, App
5	9	0.9	747	3	US-09-221-017B-1097	Sequence 1097, Ap
6	9	0.9	20284	4	US-09-526-193A-21	Sequence 21, Appl
C 7	9	0.9	53806	4	US-09-949-016-12572	Sequence 12572, A
C 8	8	0.8	255	3	US-09-134-001C-2194	Sequence 2194, Ap
C 9	8	0.8	266	4	US-09-233-785-240	Sequence 240, App
10	8	0.8	601	4	US-09-949-016-52529	Sequence 52529, A
C 11	8	0.8	601	4	US-09-949-016-52319	Sequence 52319, A
C 12	8	0.8	601	4	US-09-949-016-52320	Sequence 52320, A

; ТУРБ

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: April 17, 2005, 09:55:18 ; Search time 1145 Seconds
(without alignments)
5028.217 Million cell updates/sec
Title: US-09-671-687A-3
Perfect score: 949
Sequence: 1 MSSGLWSQKVTSPYWEERI.....RLLCAYMCMYQSPMSLYK 949
Scoring table: OLIGO 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0
Searched: 5622541 seqs, 303335566 residues
Word size: 1
Total number of hits satisfying chosen parameters: 11234615
Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Listing first 45 summaries
Command line parameters: -DEV=xl
-Q=/cgn2_1/USPTO_spool_P/US09671687/runat_15042005_163040_12059/app.query.fasta.1.1095
-DB=Published Applications NA -QMT=fastap -SUFFIX=Oligo.rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODS=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09671687@cgn.1.1.723 @runat_15042005_163040_12059
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOC=100
-LONGLOG -DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	711	74.9	4527	18	US-10-921-707-25	Sequence 25, Appl
2	685	72.2	5371	10	US-09-851-673-3	Sequence 3, Appl1
3	685	72.2	5371	18	US-10-755-889-489"	Sequence 489, App
4	641	67.5	2523	15	US-10-037-270-290	Sequence 290, App
5	641	67.5	2523	17	US-10-117-722-290	Sequence 290, App
6	279	29.4	1151	17	US-10-264-237-204	Sequence 204, App
C 7	171	18.0	617	11	US-09-969-034-4253	Sequence 4253, Ap
C 8	165	17.4	2116	18	US-10-761-370-3	Sequence 3, Appl1
C 9	148	15.6	1013	17	US-10-264-237-205	Sequence 205, App
C 10	115	12.1	557	11	US-09-969-034-4254	Sequence 4254, Ap
C 11	102	10.7	309	9	US-09-864-761-18145	Sequence 18145, A
C 12	98	10.3	483	9	US-09-864-761-1386	Sequence 1386, Ap
C 13	79	8.3	425	10	US-09-918-995-35878	Sequence 35878, A
C 14	75	7.9	376	9	US-09-983-965-1633	Sequence 1633, Ap
C 15	19	2.0	60	10	US-09-908-975-15332	Sequence 15332, A
C 16	9	0.9	201	18	US-10-719-993-10379	Sequence 10379, A
C 17	9	0.9	201	18	US-10-719-993-41266	Sequence 41266, A
C 18	9	0.9	201	19	US-10-741-600-27452	Sequence 27452, A
C 19	9	0.9	612	13	US-10-027-632-187795	Sequence 187795, A
C 20	9	0.9	612	17	US-10-027-632-187795	Sequence 187795, A
C 21	9	0.9	661	9	US-09-919-580-881	Sequence 881, App
C 22	9	0.9	842	18	US-10-425-115-90494	Sequence 90494, A
C 23	9	0.9	1440	18	US-10-437-963-63191	Sequence 63191, A
C 24	9	0.9	1863	17	US-10-424-599-31853	Sequence 31853, A
C 25	9	0.9	4294	18	US-10-437-963-63187	Sequence 63187, A
C 26	9	0.9	6574	13	US-10-194-163-1097	Sequence 1097, Ap
C 27	9	0.9	9181	10	US-09-984-827-10	Sequence 10, Appl
C 28	9	0.9	9284	17	US-10-452-510-21	Sequence 21, Appl
C 29	9	0.9	9284	17	US-10-617-334-21	Sequence 21, Appl
C 30	9	0.9	20284	18	US-10-744-465-21	Sequence 21, Appl
C 31	9	0.9	20284	18	US-10-833-679-21	Sequence 21, Appl
C 32	9	0.9	83719	18	US-10-388-838-45	Sequence 45, Appl
C 33	9	0.9	98146	18	US-10-322-281-126	Sequence 126, App
C 34	9	0.9	159138	18	US-10-719-993-6777	Sequence 6777, Ap
C 35	9	0.9	159138	19	US-10-741-600-27613	Sequence 17613, A
C 36	9	0.9	174481	19	US-10-662-613-6	Sequence 6, Appl1
C 37	9	0.9	183999	18	US-10-745-377-1	Sequence 1, Appl1
C 38	9	0.9	183999	18	US-10-872-113-1	Sequence 1, Appl1
C 39	9	0.9	303172	18	US-10-719-993-6890	Sequence 6890, Ap
C 40	8	0.8	25	19	US-10-719-900-49940	Sequence 49940, A
C 41	8	0.8	25	19	US-10-719-900-271965	Sequence 271965, A
C 42	8	0.8	25	19	US-10-719-900-520291	Sequence 520291, A
C 43	8	0.8	25	19	US-10-719-900-555961	Sequence 555961, A
C 44	8	0.8	25	19	US-10-719-900-737630	Sequence 737630, A
C 45	8	0.8	148	17	US-10-424-599-43421	Sequence 43421, A

ALIGNMENTS

RESULT 1
US-10-921-707-25
; Sequence 25, Application US/10921707
; Publication No. US20050003447A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BANDMAN, Olga
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0594 PCT
; CURRENT APPLICATION NUMBER: US/10/921,707
; CURRENT FILING DATE: 2004-08-19

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 09:17:58 ; Search time 22 Seconds
(without alignments)
4150.441 Million cell updates/sec

Title: US-09-671-687A-3
Perfect score: 949
Sequence: 1 MSSGLWSQKVTSPYWEERI.....RLLCDAYMCWQSPYMSLYK 949

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12	1.3	727	2 S42834	F40F12.5 protein -
2	12	1.3	1021	2 F88568	protein F40F12.5 [
3	8	0.8	150	2 S63996	transcription regu
4	8	0.8	215	2 C83149	probable transcrip
5	8	0.8	448	2 T15188	hypothetical prote
6	8	0.8	450	2 AG1370	glucose-6-phosphat
7	8	0.8	450	2 AE1740	protein R09B5.1 [i
8	8	0.8	458	2 D88950	yait protein precu
9	8	0.8	486	2 C64765	cytochrome P450 1A
10	8	0.8	524	2 JS0746	probable AMP-bindi
11	8	0.8	551	2 S63361	probable AMP-bindi
12	8	0.8	651	2 D83395	hydrogenase expres
13	8	0.8	740	2 E69420	ribonuclease - Dei
14	8	0.8	760	2 P75530	hypothetical prote
15	8	0.8	1086	2 T33893	hypothetical prote
16	7	0.7	50	2 A71570	hypothetical prote
17	7	0.7	72	2 G97134	hypothetical prote
18	7	0.7	81	2 C97140	hypothetical prote
19	7	0.7	104	2 T13628	hypothetical prote
20	7	0.7	116	2 I63342	tachykinin B precu
21	7	0.7	116	2 A43779	neurokinin B precu
22	7	0.7	123	2 T09268	probable tail comp
23	7	0.7	126	2 C82665	hypothetical prote
24	7	0.7	143	2 D71911	ribonuclease hi -
25	7	0.7	152	2 A10448	probable ribonucle
26	7	0.7	152	2 A84303	hypothetical prote
27	7	0.7	157	2 S58012	probable olfactory
28	7	0.7	159	2 AD1590	hypothetical prote
29	7	0.7	163	2 JC5045	epithelial membran

30	7	0.7	169	2 C91130	hypothetical prote
31	7	0.7	169	2 C85975	hypothetical prote
32	7	0.7	169	2 A65103	pts system, n-acet
33	7	0.7	171	2 E95002	conserved hypotet
34	7	0.7	174	2 A45355	neuromedin U precu
35	7	0.7	177	2 F98739	hypothetical prote
36	7	0.7	178	2 A75578	transcription regu
37	7	0.7	179	2 AF1362	hypothetical prote
38	7	0.7	180	2 A69387	fumarate (fum-1) h
39	7	0.7	189	2 F90780	probable antirepre
40	7	0.7	189	2 E90820	antirepressor prot
41	7	0.7	189	2 A90852	antirepressor prot
42	7	0.7	189	2 E90874	probable antirepre
43	7	0.7	189	2 A90902	probable antirepre
44	7	0.7	189	2 B90911	probable antirepre
45	7	0.7	189	2 G90999	probable antirepre

ALIGNMENTS

RESULT 1

S42834
F40F12.5 protein - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 24-Nov-1999
C:Accession: S42834
R:Kershaw, J.

submitted to the EMBL Data Library, February 1994

A:Reference number: S42830

A:Accession: S42834

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-727 <KER>

A:Cross-references: EMBL:Z30215

C:Genetics:

A:Introns: 19/1; 53/2; 106/3; 156/2; 198/1; 358/2; 441/2; 678/3

C:Superfamily: Caenorhabditis elegans F40F12.5 protein

Query Match 1.3%; Score 12; DB 2; Length 727;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 856 AVLCTSHYVA 867

Db 592 AVLCTSHYVA 603

RESULT 2

F88568
protein F40F12.5 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001

C:Accession: F88568

R:Anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biol.

A:Reference number: A75000; PMID:99069613; PMID:9851916

A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_e

A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; a

A:Accession: F88568

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1021 <STO>

A:Cross-references: GB:chr_III; PIDN:CAA82939.1; PID:G3877002; GSPDB:GN000021; CESP:F40

C:Genetics:

A:Gene: F40F12.5

A:Map position: 3

Query Match 1.3%; Score 12; DB 2; Length 1021;
Best Local Similarity 100.0%; Pred. No. 0.0023;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 17, 2005, 06:05:31 ; Search time 6188 Seconds
(without alignments)
5837.588 Million cell updates/sec

Title: US-09-671-687A-3
Perfect score: 949
Sequence: 1 MSSGLMSQKVTSPYWEERI.....RLLCDAYMCVQSPMTSLYK 949

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 1

Total number of hits satisfying chosen parameters: 68477535

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+p2n.model -DEV=xlp
-Q/cgn2_1/USPTO_spool_p/US09671687/runat_15042005_163038_11962/app_query.fasta_1.1095
-DB=EST -QFMT=fastap -SUFFIX=oligo.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09671687 @CEN_1_1_5180 @runat_15042005_163038_11962 -NCPU=6 -ICPU=3
-NO_WMAP -LARGOQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRAIDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	387	40.8	3496	3 AF161542	AF161542 Homo sapi
2	313	33.0	2862	3 AY406374	AY406374 Homo sapi
3	312	32.9	2862	9 AY406375	AY406375 Pan trogl
4	284	29.9	1039	4 BM457960	BM457960 AGENCOURT
5	284	25.6	1063	5 BQ433523	BQ433523 AGENCOURT
6	237	25.0	726	7 CR772310	CR772310 DKF2p468L
7	218	23.0	677	6 CD626856	CD626856 56076837J
8	208	21.9	697	4 BM724143	BM724143 UI-B-E01-
9	207	21.8	624	7 CN335014	CN335014 170005319

10	200	21.1	601	7	CR789083	CR789083 DKF2p468L
11	197	20.8	596	6	CB286853	CB286853 CMD47 E07
12	189	19.9	575	5	BP226285	BP226285 BP226285
13	189	19.9	581	7	CF528626	CF528626 UI-1-BC1-
14	185	19.5	736	6	CD641429	CD641429 AGENCOURT
15	179	18.9	919	4	BG180100	BG180100 602329675
16	178	18.8	583	5	BP219825	BP219825 BP219825
17	175	18.4	689	6	CD639981	CD639981 AGENCOURT
18	174	18.3	783	4	BM016881	BM016881 603643383
19	173	18.2	644	1	AA044899	AA044899 2k75d10.8
20	170	17.9	510	2	BF724264	BF724264 bx02g11.1
21	155	16.3	473	1	AA833748	AA833748 cd61b06.8
22	155	16.3	567	5	BP221754	BP221754 BP221754
23	155	16.3	767	6	CD370161	CD370161 UI-H-FT1-
24	154	16.2	573	5	BP221744	BP221744 BP221744
25	153	16.1	907	5	BQ438227	BQ438227 AGENCOURT
26	141	14.9	425	1	AI130924	AI130924 qc1la01.x
27	139	14.6	447	1	AA152263	AA152263 z103a04.8
28	139	14.6	473	1	AA256728	AA256728 z822h05.1
29	137	14.4	708	6	CB466941	CB466941 732655 MA
30	135	14.2	753	4	BI910868	BI910868 603067618
31	134	14.1	628	1	AA044687	AA044687 2k75d10.1
32	134	14.1	1007	5	BUS70801	BUS70801 AGENCOURT
33	133	14.0	406	1	AA761514	AA761514 rz27d01.8
34	131	13.8	541	4	BG574981	BG574981 602598070
35	131	13.8	569	1	AA056485	AA056485 2k80b08.8
36	131	13.8	635	4	BM820463	BM820463 K-EST0089
37	129	13.6	433	6	CD626855	CD626855 56076837H
38	129	13.6	565	4	BM832976	BM832976 K-EST0107
39	129	13.6	654	6	CB555752	CB555752 NMSP0069
40	129	13.6	1134	5	BM904688	BM904688 AGENCOURT
41	127	13.4	873	4	BI907475	BI907475 603063577
42	125	13.2	633	4	BG756164	BG756164 602713465
43	124	13.1	582	5	BP360904	BP360904 BP360904
44	124	13.1	837	1	AUI37267	AUI37267 AUI37267
45	123	13.0	758	1	AUI22742	AUI22742 AUI22742

ALIGNMENTS

RESULT 1
AF161542

LOCUS AF161542 Homo sapiens HSPC057 mRNA, complete cds. linear HTC 22-MAY-2001

DEFINITION AF161542.1 GI:6841351

ACCESSION AF161542

VERSION AF161542.1

KEYWORDS HTC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 3496)

AUTHORS Zhang, Q.H., Ye, M., Wu, X.Y., Ren, S.X., Zhao, M., Zhao, C.J., Fu, G.,

Shen, Y., Fan, H.Y., Lu, G., Zhong, M., Xu, X.R., Han, Z.G., Zhang, J.W.,

Tao, J., Huang, Q.H., Zhou, J., Hu, G.X., Gu, J., Chen, S.J., and Chen, Z.

Cloning and functional analysis of cDNAs with open reading frames

for 300 previously undefined genes expressed in CD34+ hematopoietic

stem/progenitor cells

Genome Res. 10 (10), 1546-1560 (2000)

JOURNAL MEDLINE 20499367

PUBMED 11042152

REFERENCE 2 (bases 1 to 3496)

AUTHORS Zhang, Q.H., Ye, M., Zhou, J., Shen, Y., Wu, X.Y., Guan, Z.Q., Wang, L.,

Fan, H.Y., Mao, Y.F., Dai, M., Huang, Q.H., Chen, S.J., and Chen, Z.

Human full length cDNA cloned from cd34+ stem cells

Unpublished

3 (bases 1 to 3496)

AUTHORS Zhang, Q.H., Ye, M., Zhou, J., Shen, Y., Wu, X.Y., Guan, Z.Q., Wang, L.,

Fan, H.Y., Mao, Y.F., Dai, M., Huang, Q.H., Chen, S.J., and Chen, Z.

Direct Submission

Submitted (14-MAY-1999) Shanghai Institute of Hematology, Shanghai

Second Medical University, Rui-Jin Hospital, 197 Rui-Jin Road II,

Shanghai 200025, P. R. China

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 09:15:43 ; Search time 65 seconds
(without alignments)

7476.356 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 949

Sequence: 1 MSSGLMSQEKVTSFYWEERI.....RLLCDAYMCYQSTMSLYK 949

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Uniprot_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	606	63.9	956	1 CYLD HUMAN	Q9nc7 homo sapien
2	116	12.2	952	1 CYLD MOUSE	Q80tg2 mus musculu
3	116	12.2	952	2 Q66h62	Q66h62 rattus norv
4	98	10.3	454	2 Q6TXJ6	Q6txj6 rattus norv
5	13	1.4	778	2 Q7Q4I4	Q7q4i4 anopheles g
6	12	1.3	550	2 Q8IPC4	Q8ipc4 drosophila
7	12	1.3	551	2 Q8IPC3	Q8ipc3 drosophila
8	12	1.3	639	2 Q8IPC5	Q8ipc5 drosophila
9	12	1.3	639	2 Q8SVF0	Q8svf0 drosophila
10	12	1.3	1144	2 Q7JMS4	Q7jms4 caenorhabdi
11	10	1.1	517	2 Q7VJ04	Q7vj04 drosophila
12	9	0.9	225	2 Q6AQ83	Q6aq83 desulfotale
13	8	0.8	80	2 Q8EX39	Q8ex39 mycoplasma
14	8	0.8	150	1 Y8BA BACSU	P17503 bacillus su
15	8	0.8	212	2 Q7QV16	Q7qv16 giardia lam
16	8	0.8	215	2 Q9HX43	Q9hx43 pseudomonas
17	8	0.8	216	2 Q9GKF5	Q9gkf5 ornithorhyn
18	8	0.8	225	2 Q7SFV2	Q7sfv2 neurospora
19	8	0.8	236	2 Q9U0S3	Q9u0s3 nilaparvata
20	8	0.8	237	2 Q8VXG1	Q8vxg1 lycopersiell
21	8	0.8	258	2 Q7RJA3	Q7rja3 anopheles g
22	8	0.8	274	2 Q6ZLGO	Q6zlg0 burkholderi
23	8	0.8	276	2 Q63WB0	Q63wb0 burkholderi
24	8	0.8	353	2 Q9EZU8	Q9ezu8 serratia ma
25	8	0.8	354	2 Q8N4A4	Q8n4a4 homo sapien
26	8	0.8	366	2 Q6AL78	Q6al78 cryptospori
27	8	0.8	371	2 Q9AWT2	Q9awt2 cryza sativ
28	8	0.8	401	2 Q9VJH4	Q9vjh4 bdellovibri
29	8	0.8	418	2 Q9VHV0	Q9vhv0 drosophila
30	8	0.8	423	2 Q01754	Q01754 caenorhabdi
31	8	0.8	450	1 G6PI_LISIN	Q928r6 listeria in

32 8 0.8 450 1 G6PI_LISMO
33 8 0.8 450 2 Q71X61
34 8 0.8 458 2 Q44617
35 8 0.8 522 2 Q6NAC1
36 8 0.8 524 1 CP11_MESAU
37 8 0.8 551 1 AG12_YEAST
38 8 0.8 555 2 Q818V0
39 8 0.8 608 1 ACEK_RALSO
40 8 0.8 625 2 Q8K7M4
41 8 0.8 651 2 Q912B2
42 8 0.8 740 2 Q28905
43 8 0.8 753 2 Q8F356
44 8 0.8 760 2 Q9RXG0
45 8 0.8 763 2 Q72SG2

ALIGNMENTS

RESULT 1

CYLD HUMAN
ID CYLD_HUMAN STANDARD; PRT; 956 AA.
AC Q9NC7; Q94934; Q7L3N6; Q96EH0; Q9NZX9;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DE Probable ubiquitin carboxyl-terminal hydrolase CYLD (EC 3.1.2.15)
DE (Ubiquitin thioesterase CYLD) (Ubiquitin-specific processing protease
DE CYLD) (Deubiquitinating enzyme CYLD) (HSPC057).
GN Names:CYLD; Synonyms:CYLD1, KIAA0849;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), DISEASE, AND TISSUE SPECIFICITY.
RX MEDLINE=20296617; PubMed=10835629; DOI=10.1038/76006;
RA Blair E., Hofmann B., Siebert R., Turner G., Evans D.G., Hansen J.,
RA Schrander-Stumpel C., Beemer F.A., Van Den Ouweland A., Halley D.,
RA Delpech B., Cleveland M.G., Leigh I., Leisti J., Rasmussen S.,
RA Wallace M.R., Fenske C., Banerjee P., Oiso N., Chaggar R., Merrett S.,
RA Leonard N., Huber M., Hohl D., Chapman P., Burn J., Swift S.,
RA Smith A., Ashworth A., Stratton M.R.,
RT Identification of the familial cylindromatosis tumor suppressor
RT Gene.;
RL Nat. Genet. 25:160-165 (2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=99156230; PubMed=10048485;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.,
RT Prediction of the coding sequences of unidentified human genes. XII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.;
RL DNA Res. 5:355-364 (1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=22388257; PubMed=1247932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R., Jordan H., Moore T., Max S.I., Wang J., Haigh P.,
RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,